

Analysis of the Convergence of the General Coupling Operator Method for One-Configuration-Type Wave Functions

JOSEP MARIA BOFILL,^{1*} HUGO BONO,² JAIME RUBIO^{2*}

¹Departament de Química Orgànica, Universitat de Barcelona, E-08028 Barcelona, Catalunya, Spain

²Departament de Química Física, Universitat de Barcelona, E-08028 Barcelona, Catalunya, Spain

Received 14 July 1997; accepted 19 September 1997

ABSTRACT: We comment on the convergence of the general coupling operator for all types of one-configuration or multiconfigurational wave functions that still preserve the one-configuration structure for the energy expression. The choice on the best arbitrary real and antisymmetric parameters inherent in the coupling operator methodology is discussed, giving a theoretical reason. Another type of coupling operator is defined, presented, and analyzed. Finally, we give some numerical examples related to the low-lying electronic states of a cluster model for K_2NiF_4 solid. © 1998 John Wiley & Sons, Inc. *J Comput Chem* 19: 368–376, 1998

Keywords: general coupling operator; coupling operator; wave functions

Introduction

One of the main problems in quantum chemistry research is the optimization of a molecular energy functional with respect to a set of orthonormal orbitals $\{\phi_m\}$. Here each c_m will denote a column that contains the coefficients of the expansion of the orbital function ϕ_m in a finite basis. This leads to the well-known self-consistent

field theory (SCF).¹ In the solution to this problem one is forced to use a variational method. With this methodology one achieves a set of orthonormal orbitals such that the energy functional is stationary—character minimum—with respect to any allowed variations in any of the orbitals of the set. Because the allowed variations should preserve the orthonormality of the orbitals set, only the orthogonal transformations U between orbitals must be considered, that is,

$$c'_i = \sum_k c_k U_{ki}. \quad (1)$$

The SCF¹ variational equations come from the energy expansion with respect to the orbital varia-

*Correspondence to: J. M. Bofill, e-mail: jmbofill@canigo.qo.ub.cs or J. Rubio, e-mail: jaime@hal6001.qf.ub.es

Contract/grant sponsor: Spanish DGICYT; contract/grant numbers: PB95-0278-C02-01, PB95-0847-C02-01

tions, taking into account the orbital orthonormality restriction. Historically, this has been defined two ways to optimize the orbital functional energy. The oldest way²⁻⁴ consists of the optimization of each orbital independently and fixation of the other occupied orbitals. Requiring the energy to be stationary under this type of variations, we get a set of eigenvalue equations

$$\mathbf{F}_i \mathbf{c}_i = \sum_j \lambda_{ij} \mathbf{S} \mathbf{c}_j \quad \forall i \quad (2a)$$

$$\lambda_{ij} = \lambda_{ji}^* \quad \forall ij \quad (2b)$$

where the set of Fock matrices $\{\mathbf{F}_i\}$ depends on all occupied orbitals and \mathbf{S} is the overlap matrix. One way to solve iteratively equations (2a) and (2b) consists of keeping the Fock matrices fixed in each iteration. In this case the set of pseudoeigenvalue equations (2) can be expressed by a coupling operator in such a way that, in its matrix representation, these equations are transformed into a single pseudoeigenvalue equation. By diagonalization of this pseudoeigenvalue equation (coupling operator method) we obtain, in principle, an improved orbitals set. Except for closed shell and some open shell cases, the coupling operator technique has a poor convergence. Many methods have been developed to improve the convergence of this algorithm. The most powerful of these improving techniques is the one proposed by Pulay in 1980, which is called direct inversion of the iterative subspace (DIIS).⁵ Briefly, the reason for the poor convergence of the coupling operator method is that equations (2a) and (2b) are the Euler-Lagrange conditions¹ of a variational principle to be satisfied by the set of the optimum orbitals. Therefore, equations (2a) and (2b) do not specify the way to modify a set of nonoptimum orbitals.

To avoid the latter problem, many authors⁶⁻¹³ have obtained variational equations that indicate how to modify a set of trial orbitals. These equations are based on the expansion of the electronic energy beyond first order in the orbital variations. The approach is based on the original idea of Levy,¹⁴ who expressed the orthogonal matrix \mathbf{U} as

$$\mathbf{U} = \mathbf{e}^{\mathbf{R}} = \mathbf{I} + \mathbf{R} + \frac{1}{2} \mathbf{R}^2 + \dots \quad (3)$$

where \mathbf{I} is the unit matrix and \mathbf{R} is an antisymmetric matrix representing a set of independent orbital rotations. Substituting equation (3) in the elec-

tronic energy expression and expanding up to second order in orbital rotations, one gets a quadratic energy approximation with respect to these parameters. The solution of this quadratic energy leads to a Newton-Raphson procedure.⁶⁻¹³ For this reason, the orbital optimization is quadratically convergent.

The latter method is quite robust and has been applied successfully to a large number of problems. Now this type of calculation is routine in quantum chemistry research to obtain a zero-order wave function. In fact, this algorithm not only solves the convergence of the one-configuration case but the more difficult multiconfiguration SCF (MCSCF) wave functions.

Owing to the complete success of the Newton-Raphson-type orbital optimization methods, the coupling operator technique has almost been abandoned. The reason for this is that the coupling operator method has been associated with the first-order orbital optimization, which has a very poor convergence. To our knowledge, during the past 20 years, only four papers have appeared intending to improve the convergence of the coupling operator method. The first three by Kosugi and Kuroda¹⁵ and Ishikawa et al.,¹⁶ incorporated in a very clever way some second-order orbital rotation through the arbitrariness of some parameters that are inherent in the nature of the coupling operator technique.^{1,4} More recently, based on the same idea of Ishikawa et al.,¹⁶ Muller et al.¹⁷ presented the same type of coupling operator method but using a different class of diagonal elements. Also Muller et al.¹⁷ used the DIIS technique⁵ to accelerate the convergence. Employing the generalized valence bond (GVB)¹⁸ wave function, Muller et al.¹⁷ showed that in this form the coupling operator method has a quite good convergence, even with this type of multiconfiguration wave function.

For a big system, let us say a large molecule and/or a large basis set, the Newton-Raphson orbital optimization requires too much computer core memory owing to the large dimension of the orbital augmented Hessian.^{11,19} In these situations, the coupling operator method presents some advantages with respect to the Newton-Raphson techniques. First of all, the inverse of the orbital Hessian is simultaneously multiplied by the orbital gradient^{16,17} and is never stored. Second, the orthogonal transformation matrix is obtained automatically from the diagonalization of the matrix representation of the coupling operator. In the Newton-Raphson-type algorithms, one has to build

the orthogonal transformation matrix from the optimized orbital rotation parameters.

In this paper, we present a revision of the coupling operator technique using different diagonal elements and starting orbitals than those used by Muller et al.¹⁷ Also, we compare the convergence of the Newton-Raphson optimization with a coupling operator that incorporates all full second-order orbital rotation. We emphasize that both Ishikawa et al.¹⁶ and Muller et al.¹⁷ techniques never use the full second-order in the construction of the coupling operator. Finally, third-order correction using the coupling operator is presented and discussed. The paper is organized in the following way: first, we present a brief discussion of the optimization of the electronic energy by orbital rotation parameters and a revision of the theory of the coupling operator. Next, comparisons and calculations with a large basis set are presented.

Theoretical Basis

EXPANSION OF THE ELECTRONIC ENERGY

Because this topic has been reviewed many times (e.g., see Bobrowicz and Goddard⁸), we give only a brief summary. For the closed and open shell SCF cases and for the GVB wave functions, that is, for an MCSCF where all Slater determinants differ from each other by at least a double replacement, the energy is given by

$$E = \sum_i^{\text{occupied orbitals}} 2v_i h_{ii} + \sum_{ij}^{\text{occupied orbitals}} (a_{ij} J_{ij} + b_{ij} K_{ij}) \quad (4)$$

where $h_{ii} = \mathbf{c}_i^T \mathbf{h} \mathbf{c}_i$; $J_{ij} = \mathbf{c}_i^T \mathbf{J}_j \mathbf{c}_i$; $K_{ij} = \mathbf{c}_i^T \mathbf{K}_j \mathbf{c}_i$ and \mathbf{h} , $\{\mathbf{J}_j\}$, and $\{\mathbf{K}_j\}$ are the one-electron, Coulomb, and exchange matrices, respectively. The set of numbers $\{v_i\}$, $\{a_{ij}\}$ and $\{b_{ij}\}$ are the so-called state parameters. Now, applying expression (3) expanded until third order in \mathbf{R} matrix into equation (4), and if only a single rotation between two orbitals, say ϕ_i and ϕ_j , is to be performed, then the energy variation is

$$\Delta E_{ij}^{(3)} = A_{ij} R_{ij} + \frac{1}{2} B_{ij,ij} (R_{ij})^2 + \frac{1}{3} C_{ij,ij,ij} (R_{ij})^3 \quad \forall i, j \quad (5)$$

where R_{ij} is the ij element of the \mathbf{R} matrix. The other matrix elements that appear in equation (5) are defined as

$$A_{ij} = \mathbf{c}_i^T (\mathbf{F}_j - \mathbf{F}_i) \mathbf{c}_j \quad (6a)$$

$$B_{ij,ij} = \mathbf{c}_i^T (\mathbf{F}_j - \mathbf{F}_i) \mathbf{c}_i - \mathbf{c}_j^T (\mathbf{F}_j - \mathbf{F}_i) \mathbf{c}_j + \gamma_{ij} \quad (6b)$$

$$\gamma_{ij} = 2(a_{ii} + a_{jj} - 2a_{ij}) K_{ij} + (b_{ii} + b_{jj} - 2b_{ij})(J_{ij} + K_{ij}) \quad (6c)$$

$$C_{ij,ij,ij} = 2\mathbf{c}_i^T (\mathbf{F}_i - \mathbf{F}_j) \mathbf{c}_j + 3(a_{ii} + a_{jj} - 2a_{ij} + b_{ii} + b_{jj} - 2b_{ij}) \times [\mathbf{c}_i^T (\mathbf{J}_i - \mathbf{J}_j) \mathbf{c}_j]. \quad (6d)$$

The set of matrices $\{\mathbf{F}_i\}$ are the so-called Fock matrices defined as

$$\mathbf{F}_i = v_i \mathbf{h} + \sum_{j \in O} (a_{ij} \mathbf{J}_j + b_{ij} \mathbf{K}_j). \quad \forall i \in O \quad (7)$$

Because $\mathbf{F}_i = \mathbf{F}_j$ for any $\forall i, j \in O$ one simplifies the notation writing \mathbf{F}_O ; that is, \mathbf{F}_O is the Fock matrix associated with the shell or subspace of orbitals O . Note that $R_{ij} = -R_{ji}$, $A_{ij} = -A_{ji}$, $B_{ij,ij} = B_{ji,ji}$ and $C_{ij,ij,ij} = -C_{ji,ji,ji}$. Optimizing $\Delta E_{ij}^{(3)}$ given in equation (5) with respect to the rotation between the orbitals ϕ_i and ϕ_j , we get⁷

$$\frac{\partial \Delta E_{ij}^{(3)}}{\partial R_{ij}} = A_{ij} + B_{ij,ij} (R_{ij}) + C_{ij,ij,ij} (R_{ij})^2 = 0. \quad (8)$$

The latter equations have been rediscovered many times, but, to our knowledge, Yaffe and Goddard⁷ were the first to give a method for the orbital optimization using equation (8). They also reported the convergence behavior of this method.

COUPLING OPERATOR METHODOLOGY AND ITS RELATION TO OTHER ORBITAL OPTIMIZATION METHODS

As was pointed out by Carbó and Riera,¹ the term *coupling operator* will be understood as a hermitic (symmetric) operator, H , constructed in such a way that the variational conditions of the electronic energy can be written as a single pseudosecular equation in Roothaan's sense:

$$\mathbf{H} \mathbf{c}'_i = \epsilon'_i \mathbf{S}'_i \quad \forall i \quad (9)$$

where \mathbf{H} is the matrix representation in the finite-basis of the operator H and

$$\mathbf{H} = \sum_{ij} \mathbf{S} \mathbf{c}_i H_{ij} \mathbf{c}_j^T \mathbf{S}. \quad (10)$$

Here, pseudosecular means that $H_{ij} = \mathbf{c}_i^T \mathbf{H} \mathbf{c}_j$, depending on $\{\mathbf{c}_i\}$. If \mathbf{H} is a hermitic (symmetric) matrix, that is, $H_{ij} = H_{ji}$, and exists as a unitary (orthogonal) matrix transformation that diagonalizes \mathbf{H} , we can find a set of vectors $\{\mathbf{c}'_i\}$ such that equation (9) is satisfied.

In the first type of coupling operator method,^{1,3,4} the elements H_{ij} of their matrix representation were chosen in the following form

$$H_{ij} = \rho_{ij}(\lambda_{ij} - \lambda_{ji}^*) = \rho_{ij} A_{ij}, \quad \forall i, j \quad (11)$$

To preserve the hermitic character of H_{ij} , the arbitrary scalars $\{\rho_{ij}\}$ have the antihermitic (antisymmetric) property $\rho_{ij} = -\rho_{ji}^*$. As was pointed out in the Introduction, this type of coupling operator method does not indicate the way to modify a set of nonoptimum orbitals.

Let us now compare the Newton-Raphson-like procedure for optimizing a set of orbitals with the coupling operator technique discussed above. In the coupling operator method, the pseudoeigenvalue problem, equation (9), is solved by the following orthogonal transformation \mathbf{U} of the representative matrix \mathbf{H} of the coupling operator

$$\mathbf{U}^T \mathbf{H} \mathbf{U} = \mathbf{e} \quad (12)$$

where \mathbf{e} is a diagonal matrix. To first-order approximation, the diagonalization of the \mathbf{H} matrix gives the orbital mixing

$$R_{ij} = \frac{H_{ij}}{(H_{jj} - H_{ii})} = \frac{\rho_{ij} A_{ij}}{(H_{jj} - H_{ii})} \quad (13)$$

where equation (11) has been used. On the other hand, neglecting the third-order term in equation (8), we get

$$R_{ij} = -\frac{A_{ij}}{B_{ij,ij}}. \quad (14)$$

Comparing equations (13) and (14), we obtain

$$\rho_{ij} = -\frac{(H_{jj} - H_{ii})}{B_{ij,ij}}. \quad (15)$$

Now, substituting equation (15) into equation (11), we obtain a new definition for the off diagonal elements of the matrix associated with the coupling operator H :

$$H_{ij} = -(H_{jj} - H_{ii}) \frac{A_{ij}}{B_{ij,ij}} \quad (16)$$

or, in full matrix form,

$$\mathbf{H} = \sum_{i \neq j} -(H_{jj} - H_{ii}) \frac{A_{ij}}{B_{ij,ij}} \mathbf{S} \mathbf{c}_i \mathbf{c}_j^T \mathbf{S} + \sum_i H_{ii} \mathbf{S} \mathbf{c}_i \mathbf{c}_i^T \mathbf{S}. \quad (17)$$

This was the basic reasoning of Ishikawa et al.,¹⁶ more recently exploited by Muller et al.¹⁷ If one diagonalizes the matrix representation of the coupling operator as defined in equation (17) to first order, one recovers equation (14), that is, the second-order convergence.

Now we consider the following type of matrix representation of the coupling operator:

$$\mathbf{H} = \sum_{i \neq j} \rho_{ij} R_{ij} \mathbf{S} \mathbf{c}_i \mathbf{c}_j^T \mathbf{S} + \sum_i H_{ii} \mathbf{S} \mathbf{c}_i \mathbf{c}_i^T \mathbf{S}. \quad (18)$$

Again, $\{\rho_{ij}\}$ is a set of antisymmetric arbitrary scalars and $\{H_{ii}\}$ an arbitrary set of diagonal elements, both sets to be characterized. From equation (8) we can obtain the best \mathbf{R} matrix according to expression

$$R_{ij} = \frac{-B_{ij,ij} \pm \sqrt{B_{ij,ij}^2 - 4A_{ij}C_{ij,ij}}}{2C_{ij,ij,ij}}. \quad \forall i, j \quad (19)$$

Note that in selecting the positive root, because A_{ij} tends to zero, R_{ij} goes to zero, which indicates that this is the desirable root to be taken. The ρ_{ij} is defined as $H_{jj} - H_{ii}$, so the off-diagonal elements of the matrix representation of the coupling operator now have the following structure

$$H_{ij} = (H_{jj} - H_{ii}) \frac{-B_{ij,ij} + \sqrt{B_{ij,ij}^2 - 4A_{ij}C_{ij,ij}}}{2C_{ij,ij,ij}} \quad \forall i, j \quad (20)$$

or in full matrix form

$$\mathbf{H} = \sum_{i \neq j} \left((H_{jj} - H_{ii}) \times \frac{-B_{ij,ij} + \sqrt{B_{ij,ij}^2 - 4A_{ij}C_{ij,ij}}}{2C_{ij,ij}} \right) \times \mathbf{S} \mathbf{c}_i \mathbf{c}_j^T \mathbf{S} + \sum_i H_{ii} \mathbf{S} \mathbf{c}_i \mathbf{c}_i^T \mathbf{S}. \quad (21)$$

Again, by diagonalizing the matrix \mathbf{H} as defined in equation (21) to first-order diagonalization, one recovers equation (19), that is, the third-order convergence.

The discussion above concerns the off-diagonal elements of the matrix representation of the coupling operator, which incorporates the proper orbital mixing for orbital pairs until third order and is scaled by the difference in the diagonal elements. In this way the appropriated mixing terms are obtained when the matrix representation of the coupling operator, \mathbf{H} , is diagonalized. Following Page and McIver,²⁰ one can improve the coupling operator technique using pseudocanonical orbitals by diagonalizing each shell individually before the construction of the \mathbf{H} matrix. With these considerations, the diagonal elements are

$$H_{ii} = \mathbf{c}_i^T \mathbf{F}_I \mathbf{c}_i, \quad \forall i \in \text{shell } I \quad (22)$$

According to the previous results, we can say that the term *coupling operator* is now an hermitic (symmetric) operator, constructed in such a way that the appropriated mixing for all possible pairs of orbitals can be written as a single pseudosecular equation. This coupling operator still has two drawbacks: First, it is assumed that the coupling between different mixings of orbital pairs is small; second, during the diagonalization of its matrix representation, \mathbf{H} , the appropriated mixing between orbitals is obtained to first order. The last point means that, far from the solution, the unitary matrix, \mathbf{U} , obtained from a Newton-Raphson procedure as described elsewhere^{7,8} will be slightly different than the \mathbf{U} matrix obtained by diagonalizing the \mathbf{H} matrix given in equation (21). To avoid this difficulty in part, the level shift or related techniques are appropriate.

Implementation, Analysis, and Discussion

The SCF calculations using variations of the matrices representation of the coupling operator described in equations (17) and (21) were done for open shell systems that have displayed pathological or slow convergence. This section is divided into two parts; the first concerns the implementation of the algorithm and the second presents the analysis of some examples.

DETAILS OF THE PROGRAM

Our SCF optimization procedure is carried out as follows.

- Step 1.** Obtain initial guess orbitals $\{\mathbf{c}_i^0\}$ by diagonalizing of the matrix representation of the Nesbet operator.²¹
- Step 2.** Construct the Fock matrices in the molecular orbital space for each shell I , $\{\mathbf{F}_I\}$, and $I = \{\text{Closed, Open, Virtual}\}$.
- Step 3.** Diagonalize each \mathbf{F}_I separately.
- Step 4.** Construct the Fock matrices $\{\mathbf{F}_I\}$ in the new basis.
- Step 5.** Compute the necessary molecular integrals.
- Step 6.** Construct the \mathbf{H} matrix corresponding to the coupling operator.
- Step 7.** Back transform the \mathbf{H} matrix to the atomic basis, labelled as \mathbf{H}' .
- Step 8.** Use the DIIS procedure⁵ to obtain an extrapolated \mathbf{H}' matrix.
- Step 9.** Diagonalize \mathbf{H}' matrix and compute the energy.
- Step 10.** Compute the density matrix and test the convergence. If the convergence is not reached, go to step 2.

Step 1 implies the construction of the matrix representation of the Nesbet operator²¹ forcing the largest fractional occupancy, that is, all orbitals occupied, with the electron occupancy set equal to $(\text{total number of electrons})/(\text{total number of orbitals})$. Because all orbitals are equally occupied in this situation, the SCF-Nesbet method converges with one iteration. Starting with the resulting orbital of the highest eigenvalue, we look for the difference

between eigenvalues of two consecutive eigenvectors such that this difference is larger according to some criteria. These two eigenvectors mark the separation between the new occupied and virtual orbital sets. Filling the occupied orbitals as $(\text{total number of electrons})/(\text{total number orbitals of the new set of occupied orbitals})$, building and solving again the SCF-Nesbet matrix, we obtain a new set of orbitals. The process is continued until we achieve the desired occupancy, and the final orbitals are taken as the guess orbitals. Sometimes it occurs that in a step of this procedure the orbital symmetry is lost. In this case, one should repeat this step but redefining the set of the occupied orbitals by adding the first orbital of the virtual set.

Step 2 involves the normal construction of the Fock matrices.

In Step 3 we diagonalize each Fock matrix separately; the resulting effect is an improved set of orbitals that neglect the intrashell interactions. Note that this diagonalization does not involve any change of the energy; the total energy is invariant under rotations between orbitals of the same shell.⁴ Sometimes it is preferable to diagonalize the matrix:

$$\begin{pmatrix} \mathbf{F}_{\text{Closed}} & \mathbf{0} & \mathbf{F}_{\text{Closed-Virtual}} \\ \mathbf{0} & \mathbf{F}_{\text{Open}} & \mathbf{0} \\ \mathbf{F}_{\text{Virtual-Closed}} & \mathbf{0} & \mathbf{F}_{\text{Virtual}} \end{pmatrix} \quad (23)$$

where, according to equation (6a), $\mathbf{F}_{\text{Closed-Virtual}} = \mathbf{c}_i^T \mathbf{F}_{\text{Closed}} \mathbf{c}_v$ and \mathbf{c}_i is an orbital that belongs to the closed shell and \mathbf{c}_v to the virtual shell.

In this case the orbitals of the closed and virtual shells are mixed, producing an energy variation. Note that the resulting orbitals present an indirect coupling between the closed and virtual shells. One can continue this procedure by building the matrix given in equation (23) with the new orbitals and diagonalize again, until the "convergence" between the closed and virtual shells is reached. However, this produces strong differences between the closed and virtual orbitals with respect to the open orbitals that are fixed. Because of this, in the present algorithm, this procedure is implemented in such a way that at the second iteration the process is stopped. This device will be referred to as SCF2.

Step 4 is the same as step 2.

Step 5 involves the calculation of the electron repulsion integrals $\{J_{ij}\}$ and $\{K_{ij}\}$ associated with the tensor $B_{ij,ij}$ (see equations 6b and 6c) and

$\mathbf{c}_i^T \mathbf{J}_i \mathbf{c}_j$ associated with the tensor $C_{ij,ij,ij}$ defined in equation (6d). The latter type of integrals do not appear in a conventional SCF program.

In Step 6 the matrix representation of the coupling operator, \mathbf{H} , is set up. If the coupling operator is of second-order type, hereafter denoted SOCO (second-order coupling operator), we use equation (17), and, if it is of third order, hereafter denoted TOCO (third-order coupling operator), we employ equation (21).

In Step 8 an acceleration technique is used. The best seems to be the DIIS. Note that the DIIS technique requires the matrix representation of the coupling operator, \mathbf{H} , in atomic basis, \mathbf{H}' . In the present case, the error vector, denoted **error**, at the iteration, say, j , is

$$\mathbf{H}'^{(j)} \mathbf{D}^{(j)} \mathbf{S} - \mathbf{S} \mathbf{D}^{(j)} \mathbf{H}'^{(j)} = \mathbf{error}^{(j)} \quad (24)$$

where \mathbf{S} is the overlap integrals matrix and $\mathbf{D}^{(j)}$ is the first-order density matrix at iteration j . The DIIS technique starts when $[(\mathbf{error}^{(j)})^T (\mathbf{error}^{(j)})]^{1/2} / N \leq \varepsilon$, where N is the length of the vector $\mathbf{error}^{(j)}$. The parameter ε should be a small number ($\varepsilon = 10^{-7}$).

Step 9 involves a normal matrix diagonalization.

Finally, at Step 10, the new orbitals and density matrix are computed and the convergence criteria are also tested. The convergence criteria are taken on the differences in both the energy and the wave function. The later means that $\max |\mathbf{D}_{\mu\nu}^{(k)} - \mathbf{D}_{\mu\nu}^{(k+1)}| \leq \text{threshold}$, where the pair $\mu\nu$ runs over the basis set functions. This criterion incorporates the convergence in both the molecular orbitals and the GVB coefficients if the calculation implies the optimization of a GVB wave function.

STUDY OF A CLUSTER MODEL FOR K_2NiF_4 SOLID

As an example, we have chosen a cluster model for studying the magnetic coupling in K_2NiF_4 solid.²² The cluster possesses D_{2h} symmetry, and it has a central Ni_2F_{11} unit, where the 11 F^- are simulated by point charges embedded by 16 K^+ and six Ni^{2+} , as total ion potentials,²³ plus a total of 1,154 point charges.

For the two Ni^{2+} treated explicitly, we use the all electron basis set consisting of (14s, 11p, 6d)²⁴ contracted to (6s, 5p, 4d) as in Graaf et al.²⁵ The whole system has a total of 82 basis functions.

This example is representative of the conver-

gence difficulties that one finds in this type of system, as well as in the study of chemisorption processes on metal surfaces.^{26–28} We have studied the two more important parameters that strongly affect the convergence of the SCF process, namely, the guess orbitals and the order of the coupling operator (SOCO or TOCO). However, to obtain a method more efficient from a computational point of view, we have analyzed other factors important in the convergence behavior and efficiency of the present algorithm: the starting iteration of the DIIS acceleration technique⁵; the use of exact or approximated Coulomb and Exchange integrals, J_{ij} and K_{ij} , respectively; and finally the SCF2 method proposed in the previous section.

First we note that in the computation of transition metal clusters, convergence problems appear mainly because there exist many electronic states almost energetically degenerated, which have a quite different electron distribution. These states come from different *d*-electron distributions. We show in Table I that the present example falls into this case.

In the first series of calculations the starting orbitals were Nesbet-type orbitals,²¹ with 45 orbitals with an occupancy of 1.15 electrons each, hereafter labelled as OM1. These orbitals still present an occupancy that is far from the real situation, that is, 24 molecular orbitals doubly occupied and four orbitals with one electron. The results are summarized in Table II. First, we observe that the convergence depends strongly on whether one uses SOCO or TOCO. The TOCO method always converges in a low number of iterations. However, we observe that the behavior of the two methods is different. The difference depends on the starting point of the DIIS technique and the type of integrals used. If the exact integrals are used to build

the matrix associated with the operator, the TOCO always converge to the same state, 1^5B_{3g} independently of the threshold associated with the DIIS technique, whereas the SOCO converges to three different electronic states, none of them corresponding to the ground state.

These results are explained by taking into account that in the TOCO method, when using the exact integrals, the convergence radius is determined by the order of the Taylor series expansion of the rotation orbitals. In this situation nothing indicates the correct search direction. In other words, at third-order expansion, the algorithm finds a radius of convergence that is as close as possible to the initial value and prevents any possible redistribution of the occupied orbitals. In contrast, with the coupling operator method that uses the rotation energy expansion to second order, SOCO, the occupancy of the orbitals changes very much during the optimization process and the final converged electronic state depends on the starting DIIS criterion. In this case, the radius of convergence is much smaller than with TOCO and does not exist at any stationary point within this domain; consequently, the algorithm involves an electronic distribution until a stationary point is found.

When the integrals $\{J_{ij}\}$ and $\{K_{ij}\}$ that appear in equations (6c) and (6d) are computed approximately rather than exactly, the SOCO is not quite as sensitive to the starting DIIS criterion and the process converges to one of the electronic states found when the exact integrals are used. On the other hand, TOCO method with approximated integrals presents an increase of the convergence radius and, consequently, during the optimization process the orbitals are mixed much more than when the integrals are calculated in an exact way.

TABLE I.
Electronic Configuration, Symmetry, and Energy of the First Six Electronic States of K_2NiF_4 .

Configuration	State symmetry	Energy ^a	Δ^b
$a_g a_g b_{2u} b_{2u}$	1^5A_g	−3077.3385	0.0
$a_g b_{2u} a_u b_{2g}$	2^5A_g	−3077.3232	0.0153 (9.6)
$a_g b_{2u} b_{1g} b_{3u}$	3^5A_g	−3077.3224	0.0161 (10.1)
$b_{1g} b_{3u} b_{1u} b_{3g}$	4^5A_g	−3077.2752	0.0633 (30.7)
$a_u b_{2g} b_{1u} b_{3g}$	5^5A_g	−3077.2711	0.0674 (42.3)
$a_g a_u b_{2g} b_{1u}$	1^5B_{3g}	−3076.8988	0.4397 (275.8)

^aThe energy is given in Hartree.
^bThe energy difference with respect to the 1^5A_g state is given in Hartree. In parenthesis, this difference is given in kcal/mol.

TABLE II.
Convergence Behavior of the Second- and Third-Order Coupling Operator Methods with the OM1 Starting Orbitals.

	Integrals ^a					
	Approximated			Exact		
Third-order coupling operator (TOCO)						
DIIS ^b	0.5	0.1	0.01	0.5	0.1	0.01
State ^c	4 ⁵ A _g	1 ⁵ A _g	1 ⁵ A _g	1 ⁵ B _{3g}	1 ⁵ B _{3g}	1 ⁵ B _{3g}
Iterations	11	14	18	11	17	17
Second-order coupling operator (SOCO)						
DIIS ^b	0.5	0.1	0.01	0.5	0.1	0.01
State ^c	3 ⁵ A _g	3 ⁵ A _g	3 ⁵ A _g	5 ⁵ A _g	3 ⁵ A _g	2 ⁵ A _g
Iterations	17	17	20	12	26	22

^aIn SOCO, *approximated* integrals means that all the integrals that appear in the matrix elements γ_{ij} , equation (6c), are set equal to one; in TOCO, it means that all integrals that appear in the matrix elements γ_{ij} are set equal to one, but the integrals $\mathbf{c}_i^T \mathbf{J}_i \mathbf{c}_j = \langle ii | ij \rangle$ that appear in the tensor elements $C_{ij,ij,ij}$, equation (6d), are calculated exactly. *Exact* means that all these integrals are computed correctly.

^bThe starting DIIS technique threshold.

^cThe symmetry of the final converged electronic state.

Now, the method is able to converge to the absolute minima in a reduced number of iterations.

The explanation for this fact is given by considering that higher orders in the expansion of the energy have been omitted, and this is compensated for by the use of approximated integrals. Of course, one can change the values of the approximated integrals; however, using these values in other systems, the process converges very well.^{26–28}

Regarding the SCF2 method, starting at OM1 orbitals that are far from the optimum orbitals of the ground state, the SOCO (TOCO) method always converges to a state that is 29.3 (29.4) au above the ground state in eight iterations. That is, the SCF2 method converges after a few iterations to the closer state of the initial orbitals.

Now, we will study the influence of the initial orbitals on the convergence. Because the initial orbitals determine the convergence process, we have employed other guess orbitals obtained through the Nesbet procedure but with an occupancy of 1.4 electrons in 36 molecular orbitals (OM2). This occupancy is closer to the final one. In this case all the coupling operators proposed converge to the same excited state 4⁵A_g, and depending on the different options the number of iterations ranges from 12 to 17. This result turned out to be quite obvious; an analysis of the Nesbet orbitals showed that they are very close to the orbitals of the 4⁵A_g state. This explains the fact that both second and third order converge to the

same state in this case. However, the proposed SCF2 method is more efficient, insofar as the radius of convergence is reached just when the iterative process starts and converges after few iterations to the ground state.

Conclusions

We have presented a study of the convergence of the coupling operator technique for solving the SCF equations. Using decoupled orbital pair rotations, we have defined a new coupling operator, which involves mixings between orbitals to second and third order in the energy expansion. This technique has been applied to the computation of the first six electronic states of a K₂NiF₄ cluster. The main conclusion is that the TOCO method converges very fast, and the SOCO method converges with a few more iterations, but both always converge. Reordering these converged orbitals and using them as starting orbitals, we can easily obtain the desired ground state.

Acknowledgments

We are indebted to Profs. F. Illas and S. Olivella for their valuable suggestions. We thank the “Centre de Supercomputació de Catalunya” (CESCA) for CPU time allocation.

References

1. R. Carbó and J. M. Riera, *A General SCF Theory*, Lecture Notes in Chemistry, Vol. 5, Springer Verlag, New York, 1976 (and references therein).
2. C. C. J. Roothaan, *Rev. Mod. Phys.*, **32**, 179 (1960).
3. R. McWeeny, *Mol. Phys.*, **28**, 1273 (1974); *Chem. Phys. Lett.*, **35**, 13 (1975).
4. K. Hirao and K. Nakatsuji, *J. Chem. Phys.*, **59**, 1457 (1973); K. Hirao, *J. Chem. Phys.*, **60**, 3215 (1974).
5. P. Pulay, *Chem. Phys. Lett.*, **73**, 393 (1980); *J. Comput. Chem.*, **3**, 556 (1982).
6. S. Polezzo, *Theor. Chim. Acta*, **40**, 245 (1975); P. Fantucci and S. Polezzo, *Theor. Chim. Acta*, **45**, 317 (1977).
7. L. G. Yaffe and W. A. Goddard III, *Phys. Rev. A*, **13**, 1682 (1976).
8. F. W. Bobrowicz and W. A. Goddard III, in: *Modern Theoretical Chemistry*, ed. H. F. Schaefer, Vol. III, Plenum, New York, 1977, p 79.
9. J. Kendrick and I. H. Hillier, *Chem. Phys. Lett.*, **41**, 283 (1976).
10. E. Dalgaard and P. Jørgensen, *J. Chem. Phys.*, **69**, 3833 (1978).
11. B. H. Lengsfeld III, *J. Chem. Phys.*, **73**, 382 (1980).
12. H. J. Werner and W. Meyer, *J. Chem. Phys.*, **73**, 2342 (1980).
13. H. J. Werner, *Adv. Chem. Phys.*, **69**, 1 (1987).
14. B. Levy, *Chem. Phys. Lett.*, **4**, 17 (1969); *Int. J. Quant. Chem.*, **4**, 297 (1970).
15. N. Kosugi and H. Kuroda, *Chem. Phys. Lett.*, **74**, 490 (1980).
16. Y. Ishikawa and G. L. Malli, *Chem. Phys. Lett.*, **76**, 565 (1980); Y. Ishikawa, R. C. Binning, Jr., and W. H. Adams, *J. Chem. Phys.*, **76**, 4105 (1982).
17. R. P. Muller, J.-M. Langlois, M. N. Ringnalda, R. A. Friesner, and W. A. Goddard III, *J. Chem. Phys.*, **100**, 1226 (1994).
18. W. A. Goddard III, *Science*, **227**, 917 (1985).
19. R. Shepard, *Adv. Chem. Phys.*, **69**, 63 (1987).
20. M. Page and J. W. McIver Jr., *J. Chem. Phys.*, **79**, 4985 (1983).
21. R. K. Nesbet, *Rev. Mod. Phys.*, **35**, 532 (1963).
22. P. T. Moreira and F. Illas, *Phys. Rev. B*, **55**, 1 (1997).
23. N. W. Winter and R. M. Pitzer, *J. Chem. Phys.*, **89**, 446 (1988).
24. C. W. Bauschlicher and P. S. Bagus, *J. Chem. Phys.*, **81**, 5889 (1984).
25. C. Graaf, F. Illas, R. Broer, and C. Nieuwpoort, *J. Chem. Phys.*, **106**, 3287 (1997).
26. F. Illas, J. Rubio, J. M. Ricart, and G. Pacchioni, *J. Chem. Phys.*, **105**, 7192 (1996).
27. J. Casanovas, J. Rubio, and F. Illas, *Phys. Rev. B*, **53**, 945 (1996).
28. M. Mödl, A. Povill, J. Rubio, and F. Illas, *J. Phys. Chem.*, **101**, 1526 (1997).